

BIBLIOTECA DI SCIENZE STATISTICHE

SERVIZIO BIBLIOTECARIO NAZIONALE

SID PV0850043 DID

ACQ. 478 / 102 INV. 82178

COLL. CLASS. 5-GR WP. 2002/11

**Second-Order Likelihood-
Based Inference in Mixed-
Effects Models**

A. Guolo

2002.11

**Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti 241-243
35121 Padova**

Agosto 2002

Second-Order Linear
Based Interference in Mixed-
Effects Models

A. Ghosh

1993

Department of Statistics
University of Southern California
516 C. Hallman 241-243
38121 Pasadena

Second-Order Likelihood-Based Inference in Mixed-Effects Models

Annamaria Guolo

August 2002

Abstract

This paper focuses on the application of higher-order asymptotics for likelihood-based inference to mixed-effects models. In particular, for a scalar parameter of interest, the modified directed likelihood of Skovgaard (1996) is considered. Its expression in linear and nonlinear mixed-effects models with a single level of grouping is derived, starting from the results of Lyons and Peters (2000). Some simulation studies demonstrate the improvements that can be obtained by using this statistic with respect to first-order methods.

Keywords: higher-order asymptotics, likelihood-based inference, Lindstrom and Bates' approximation, linear and nonlinear mixed-effects models, Skovgaard's statistic

1 Introduction

In economical, social and biological sciences, the analysis of grouped data is very common. Data structures with categorized observations are routinely handled by approaches like the covariance analysis. The base assumption of this method is an *a priori* definition of a finite number of categories and the representation of each of these categories in the sample by at least one observation. In this type of models, the parameters associated to the covariates are the fixed effects. Often, it is reasonable to assume that observation groups are not necessarily priorly defined, but they can be just a sample from an infinite population of groups. Therefore, inference interest focuses on the entire population of groups and not only on the differences among really observed and sampled groups. In this case, covariates parameters are realizations of random variables, called the random effects. Models with both fixed effects and random effects are called mixed-effects models.

Likelihood-based inference in mixed-effects models is typically based on first-order results. The application of higher-order asymptotics is relatively recent. Lyons and Peters (2000) derived

Skovgaard's second-order statistic $\tilde{r}$ for linear mixed-effects models. The aim of this paper is to present the derivation and the properties of $\tilde{r}$ in two main cases. First of all we consider its application to linear mixed-effects models with either homoscedastic and independent or heteroscedastic and correlated errors. Then, we extend its application to the class of nonlinear mixed-effects models.

The theoretical background of higher-order asymptotics and of linear and nonlinear mixed-effects models is briefly reviewed in section 2. For a comprehensive introduction to higher-order asymptotics, see Severini (2000). Section 3 concerns the application of Skovgaard's statistic to the above-mentioned classes of models. We report also some simulation studies that compare the second-order statistic to the first-order one, by evaluating their agreement with the standard normal asymptotic null distribution.

2 Basic techniques

2.1 Higher-order asymptotics

We consider a parametric statistical model with model function $p(y; \theta)$. Let θ be the parameter vector and $\ell(\theta)$ the log-likelihood. Suppose $\theta = (\psi, \lambda)$ is partitioned into a parameter of interest ψ and a nuisance parameter λ . Let $\hat{\theta} = (\hat{\psi}, \hat{\lambda})$ be the maximum likelihood estimate (MLE) of θ and $\tilde{\theta} = (\psi, \tilde{\lambda}_\psi)$ be the constrained MLE of θ for a fixed value of ψ .

For scalar ψ , the reference statistic for likelihood inference is the directed log-likelihood ratio

$$r(\psi) = \text{sgn}(\hat{\psi} - \psi) \sqrt{\{\ell(\hat{\theta}) - \ell(\tilde{\theta})\}}$$

which has a null standard normal distribution up to an error of order $O(n^{-1/2})$. Barndorff-Nielsen (1986) proposed a modified version of the previous statistic, indicated as $r^*(\psi)$. It presents a standard normal distribution up to an error of order $O(n^{-3/2})$, so improving the accuracy of the approximation. The $r^*(\psi)$ statistic is defined as

$$r^*(\psi) = r(\psi) - \frac{1}{r(\psi)} \log \frac{r(\psi)}{u(\psi)}. \quad (1)$$

In the (1), the correction term $u(\psi)$ involves certain partial derivatives of the log-likelihood with respect to components of the MLE (sample space derivatives) in which a complementary ancillary statistic is held fixed. They are often difficult to calculate outside full exponential and group families.

Skovgaard (1996) proposed an approximation to Barndorff-Nielsen's statistic, here denoted by $\tilde{r}(\psi)$. It is simpler in calculation, though it maintains a bigger approximation accuracy than the corresponding first-order statistic $r(\psi)$. Skovgaard's statistic $\tilde{r}(\psi)$ substitutes sample space derivatives with expected values of likelihood quantities and does not require the specification

of an ancillary statistic. Its expression is given by

$$\tilde{r}(\psi) = r(\psi) - \frac{1}{r(\psi)} \log \frac{r(\psi)}{\tilde{u}(\psi)}, \quad (2)$$

and it has a standard normal null distribution up to an error of order $O(n^{-1})$. The quantity $\tilde{u}(\psi)$ in the (2) is defined by

$$\tilde{u}(\psi) = [S^{-1}q]_{\psi} |\hat{j}|^{1/2} |\hat{i}|^{-1} |S| |\hat{j}_{\lambda\lambda}|^{-1/2},$$

where i is the expected information matrix, j is the observed information matrix, S and q are defined as follows. Let ℓ_{θ} denote the score function, then

$$S = \text{Cov}_{\hat{\theta}} \{ \ell_{\theta}(\theta_1), \ell_{\theta}(\theta_2) \} |_{\theta_1=\hat{\theta}, \theta_2=\hat{\theta}}$$

and

$$q = \text{Cov}_{\hat{\theta}} \{ \ell_{\theta}(\theta_1), \ell(\theta_1) - \ell(\theta_2) \} |_{\theta_1=\hat{\theta}, \theta_2=\hat{\theta}}.$$

The statistic (2) is relatively simple to calculate. Elements in $\tilde{u}(\psi)$ have the computational complexity of the expected information matrix. Moreover, $\tilde{r}(\psi)$ involves only likelihood quantities and so is well defined for all sufficiently regular parametric models. Finally, $\tilde{r}(\psi)$ is invariant with respect to interest respecting reparameterizations.

2.2 Mixed-effects models

2.2.1 Linear mixed-effects model

The linear mixed-effects (LME) model (Laird and Ware, 1982) considers m subjects or, more generally, groups, for each of which there are n_i observation of the response variable y_i and of the covariates x_i , for $i = 1, \dots, m$.

Let β be the p -dimensional vector of fixed effects and let b_i be the k -dimensional vector of random effects. The LME model is defined as

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i, \quad i = 1, \dots, m, \quad (3)$$

where X_i (of size $n_i \times p$) and Z_i (of size $n_i \times k$) are known fixed-effects and random-effects regressor matrices.

The LME model is specified further by the following hypotheses:

1. the within-group errors ε_i follow a normal distribution with mean zero and covariance matrix Λ_i ,

$$\varepsilon_i \sim N_{n_i}(0, \Lambda_i);$$

2. the random effects b_i follow a normal distribution with mean zero and covariance matrix Ψ ,

$$b_i \sim N_k(0, \Psi).$$

3. the random effects b_i and the within-groups errors ε_i are independent for different groups, $i = 1, \dots, m$, and (b_i, ε_i) are multivariate independent random variables of each other for the same group.

Because in (3) the response variables are modelled as the sum of two independent multivariate normal random vectors, it follows that, marginally, the y_i are independently distributed as multivariate normal vectors with mean $X_i\beta$ and covariance matrix V_i

$$Y_i \sim N_{n_i}(X_i\beta, V_i) \text{ , where } V_i = Z_i\Psi Z_i^\top + \Lambda_i \text{ .} \quad (4)$$

The covariance matrix V_i depends on a vector of variance parameters, ρ , so we can write $V_i = V_i(\rho)$. As we can see from (4), the response variability is due to two components: the within-group error vector ε_i and the random-effects vector b_i .

The parametric vector θ (cf. §2.1) is specified as

$$\theta = (\beta, \rho)$$

and it belongs to the parametric space $\Theta \subseteq \mathbb{R}^{p+\bar{k}}$, being $\bar{k}$ the dimension of ρ .

The contribution of the vector y_i to the likelihood for each of the m groups of observations is

$$\begin{aligned} L(\beta, \rho; y_i) &= p(y_i; \beta, \rho) \\ &= |2\pi V_i|^{-1/2} \exp \left\{ -\frac{1}{2} (y_i - X_i\beta)^\top V_i^{-1} (y_i - X_i\beta) \right\} . \end{aligned}$$

The likelihood for the total of observations is hence

$$\begin{aligned} L(\beta, \rho; y) &= \prod_{i=1}^m p(y_i; \beta, \rho) \\ &= \prod_{i=1}^m |2\pi V_i|^{-1/2} \exp \left\{ -\frac{1}{2} (y_i - X_i\beta)^\top V_i^{-1} (y_i - X_i\beta) \right\} \end{aligned}$$

and the associated log-likelihood is

$$\begin{aligned} \ell(\beta, \rho; y) &\propto \sum_{i=1}^m -\frac{1}{2} \log |V_i| - \frac{1}{2} \sum_{i=1}^m (y_i - X_i\beta)^\top V_i^{-1} (y_i - X_i\beta) \\ &= -\frac{1}{2} \sum_{i=1}^m \log |V_i| - \frac{1}{2} \sum_{i=1}^m \left\{ y_i^\top V_i^{-1} y_i - (X_i\beta)^\top V_i^{-1} y_i \right. \\ &\quad \left. - y_i^\top V_i^{-1} X_i\beta + (X_i\beta)^\top V_i^{-1} X_i\beta \right\} \\ &= -\frac{1}{2} \sum_{i=1}^m \log |V_i| - \frac{1}{2} \sum_{i=1}^m \left\{ y_i^\top V_i^{-1} y_i - 2(X_i\beta)^\top V_i^{-1} y_i + (X_i\beta)^\top V_i^{-1} X_i\beta \right\} . \quad (5) \end{aligned}$$

Maximum likelihood estimates for β and ρ are obtained by maximization of (5) using numeric methods, as the EM algorithm or the Newton-Raphson method (Searle *et al.*, 1992, chapter 8).

Newton-Raphson method implementation for mixed-effects models is discussed in Jennrich and Schluter (1986) and in Lindstrom and Bates (1988), which also presents a comparison with the EM algorithm. A detailed description of random effects predictions (BLUP) is given in Robinson (1991).

2.2.2 Nonlinear mixed-effects model

The nonlinear mixed-effects (NLME) model extends the LME model by allowing the regression function to depend nonlinearly on fixed and random effects. Suppose the relation between y and x is represented by a function f , which is nonlinear in the parameters, real-valued and differentiable (Lindstrom and Bates, 1990). Even though its form is common to all the m groups of observations, its parameters can vary among groups. Assumptions on random effects and within-group error distributions are the same as in the linear case.

The NLME model definition is in two steps.

1. The j th observation on the i th group is modeled as

$$y_{ij} = f(\phi_{ij}, x_{ij}) + \varepsilon_{ij} \quad , \quad i = 1, \dots, m \quad , \quad j = 1, \dots, n_i \quad , \quad (6)$$

where f is nonlinear in at least one of the components of the group specific parameter vector ϕ_{ij} , whose specification constitutes the second step.

2. The ϕ_{ij} vector is modeled as a linear combination of β and b_i , by the known matrices A_{ij} and B_{ij}

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i \quad . \quad (7)$$

We can write (6) and (7) in matrix form as

$$y_i = f_i(\phi_i(\beta, b_i)) + \varepsilon_i \quad , \quad (8)$$

where $f_i(\phi_i(\beta, b_i))$ is a reduced form for $f_i(\phi_i(\beta, b_i), x_i)$ and $\phi_i = A_i\beta + B_ib_i$, for $i = 1, \dots, m$.

Generally, the NLME model is more parsimonious than a LME model in the number of parameters required to obtain a good description of data. This advantage of NLME models does not come without costs, however. Because random effects are allowed to enter the model nonlinearly, the likelihood function does not have a closed-form expression. As a consequence, an approximate likelihood function needs to be used for inferential procedures, clearly leading to more computationally intensive estimation algorithms and to less reliable inference results. In this context, we refer to Lindstrom and Bates' (1990) approximation. For a detailed description of what has been proposed in literature, see Davidian and Giltinan (1995).

2.3 Lindstrom and Bates' approximation

Lindstrom and Bates' approximation is a linearization of model (8) in the random effects b_i . It is obtained by applying a Taylor series expansion of f_i around $b_i = b_i^*$, where b_i^* is a value which corresponds to the BLUP in the linear case. The approximation is

$$y_i \approx f_i(\phi_i(\beta, b_i))|_{b_i=b_i^*} + Z_i(\phi_i(\beta, b_i))|_{b_i=b_i^*} (b_i - b_i^*) + \varepsilon_i, \quad (9)$$

where Z_i is the matrix (of size $n_i \times k$) of the derivatives of f_i with respect to the random effects b_i . Hence, the response variable follows a normal distribution with mean and covariance matrix approximated, respectively, by

$$E(Y_i) \approx f_i(\phi_i(\beta, b_i^*)) - Z_i(\phi_i(\beta, b_i^*)) = f_i^*(\beta, b_i^*) = f_i^*$$

$$\text{Cov}(Y_i) \approx \Lambda_i + Z_i(\phi_i(\beta, b_i^*)) \Psi Z_i(\phi_i(\beta, b_i^*))^\top = V_i(\beta, b_i^*, \rho) = V_i.$$

For an exhaustive explanation of the estimation procedure for the approximate model, see Lindstrom and Bates (1990) and Davidian and Giltinan (1995, chapter 6).

3 Higher-order asymptotics for mixed-effects models

Application of higher-order asymptotics to mixed-effects models starts with Lyons and Peters (2000). They calculate Skovgaard's statistic $\tilde{r}$ in linear mixed-effects models and report some simulation studies. These clearly show the improvement in the null approximation to the standard normal with respect to the first-order method r . Lyons and Peters (2000) primarily concentrate on models with independent and identically distributed (IID) within-group errors. This assumption simplifies the calculation of all components of $\tilde{r}$, especially in the differentiation of likelihood quantities with respect to the parameter vector θ . In fact, in this case we can write the covariance matrix V_i as a linear combination of known matrices with weights given by the components of ρ , as follows

$$V_i = \sum_{r=1}^{\bar{k}} \rho_r Q_{ir}, \quad (10)$$

with known symmetric matrices $Q_{ir}, i = 1, \dots, m, r = 1, \dots, \bar{k}$, being $\bar{k}$ the dimension of ρ . Hence, we easily obtain the derivatives of V_i with respect to the components of ρ

$$V_{i\rho_r} = \frac{\partial}{\partial \rho_r} V_i = Q_{ir}.$$

Starting from Lyons and Peters' results, we consider here the extension to linear mixed-effects models in which the within-group errors ε_i can have a generic covariance structure, not necessarily IID (§3.1). Moreover, we consider the application to nonlinear mixed-effects models (§3.2).

3.1 Application to linear mixed-effects models

Consider the LME model according to the formulation given in (3). We do not report here the specific calculation of all components of $\tilde{r}$, for which we refer to Guolo (2001, chapter 2). In the following subsection we report the results of a simulation study.

3.1.1 Simulation study

We consider a linear mixed-effects model with a random effect associated with the intercept

$$y_i = \beta_0 + \beta_1 x_i + b_{i1} + \varepsilon_i, \quad (11)$$

where $b_i \sim N_1(0, \Psi)$ and $\varepsilon_i \sim N_{n_i}(0, \Lambda_i)$. The covariance matrix for within-group errors ε_i is

$$\Lambda_i = \sigma_e^2 \begin{bmatrix} \exp\{2\delta v_{i1}\} & 0 & \dots & 0 \\ 0 & \exp\{2\delta v_{i2}\} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \exp\{2\delta v_{in_i}\} \end{bmatrix}. \quad (12)$$

The covariance matrix represented by (12) is an exponential function of the variance covariate v_{ij} . The parameter δ is unrestricted, so we can model cases where the variance increases and decreases with the variance covariate v_{ij} .

The simulation study considers $m = 5$ groups with 15, 14, 15, 15, 16 observations respectively and

$$\beta = (\beta_0, \beta_1)^T = (0, 1)^T, \quad \delta = 0.5.$$

The random effect b_i follows the standard normal distribution. The covariate x_{ij} values are generated from an exponential variable with mean 2 and are maintained constant during the simulation. We let $v_{ij} = x_{ij}$. Interest focuses on $\beta_0 = 0$. Figure 1 reports the normal Q-Q plots of r and $\tilde{r}$ obtained from a simulation of $R = 1000$ replicates.

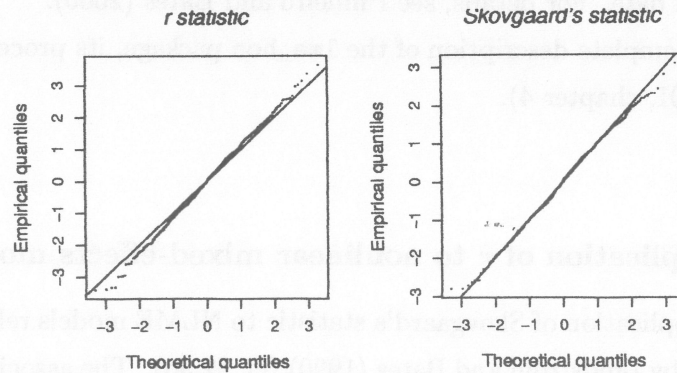


Figure 1: Normal Q-Q plots for r (left) and $\tilde{r}$ (right) based upon $R = 1000$ replicates of a simulation for testing $\beta_0 = 0$ in model (11).

The normal approximation does not work very well for r , especially in the tails of the distribution. Using the second-order statistic $\tilde{r}$, instead, we obtain a more satisfactory behaviour. A similar judgement applies to the results in Table 1, that reports the estimated coverage of confidence intervals with different sizes. Here, the tendency for r is to underestimate the nominal level, while the reduction of differences between nominal and empirical levels is substantial for $\tilde{r}$.

	Statistic	
Coverage	r	$\tilde{r}$
0.80	0.741 (0.715,0.767)	0.793 (0.767,0.819)
0.85	0.796 (0.770,0.822)	0.845 (0.823,0.867)
0.90	0.858 (0.836,0.880)	0.905 (0.887,0.923)
0.95	0.924 (0.908,0.940)	0.950 (0.936,0.964)
0.99	0.980 (0.972,0.988)	0.985 (0.977,0.993)

Table 1: Estimated coverage of 80%, 85%, 90%, 95% and 99% confidence intervals for the fixed effect β_0 in model (11) obtained from $R = 1000$ replicates of a simulation.

3.1.2 Software: the `lme_hoa` package

The simulation studies we realized for LME models have been obtained by means of a set of procedures, called `lme_hoa` package. The software has been developed in R version 1.3.0 for Microsoft Windows and it is applicable also using the successive versions.

The `lme_hoa` package allows us to calculate Skovgaard's statistic $\tilde{r}$ in linear mixed-effects models either if the ε_i are IID or if they present heteroscedasticity or correlation. For the estimation of parameters the `lme_hoa` package uses the `nlme` library of Pinheiro and Bates (2000). The `nlme` library has been constructed for R (version 1.3.0) for Microsoft Windows. It provides instruments for application of linear and nonlinear mixed-effects models in the analysis of grouped data. For details, see Pinheiro and Bates (2000).

For a complete description of the `lme_hoa` package, its procedures and their application, see Guolo (2001, chapter 4).

3.2 Application of $\tilde{r}$ to nonlinear mixed-effects models

The application of Skovgaard's statistic to NLME models relies on the model approximation suggested by Lindstrom and Bates (1990) (cf. §2.3q). The associated approximate log-likelihood

is equal to

$$\ell(\beta, \rho) = \sum_{i=1}^m \left\{ \log |V_i| + [y_i - f_i^*(\phi_i(\beta, b_i^*))]^\top V_i^{-1} [y_i - f_i^*(\phi_i(\beta, b_i^*))] \right\}, \quad (13)$$

where

$$V_i = V_i(\beta, b_i^*, \rho) = \Lambda_i + Z_i(\phi_i(\beta, b_i^*)) \Psi Z_i(\phi_i(\beta, b_i^*))^\top. \quad (14)$$

We have derived $\tilde{r}$ in NLME models with a procedure similar to the linear case. However, some differences existing between LME and NLME models have to be considered:

- the function f_i^* is nonlinear in the parameters;
- the covariance matrix V_i depends not only on ρ but also on the fixed effects β , that appear inside the random-effects matrix Z_i .

These elements make $\tilde{r}$ calculation much more complicated, most of all in differentiation of V_i with respect to both fixed effects β and to variance parameters ρ . The problem is that we do not have the separation between the fixed-effects part and the variance part of the log-likelihood, which is typical of the linear case. Another disadvantage is the impossibility to refer to a simplified form such as (10) for the covariance matrix of Y_i .

To solve some of the above computational problems, we proposed an approximation to Skovgaard's statistic. This approximation is indicated as $\tilde{r}_A$. The idea is to replace the unknown β in the covariance matrix V_i with its MLE. So we can write (14) as

$$\bar{V}_i = \bar{V}_i(\rho) = V_i(\hat{\beta}, b_i^*, \rho) = \Lambda_i + Z_i(\hat{\beta}, b_i^*) \Psi Z_i(\hat{\beta}, b_i^*)^\top$$

that now depends only on the variance parameters ρ . In this way, we re-establish the separation between the fixed-effects part and the variance part of the log-likelihood.

We realized some simulation studies to compare the behaviour of r , $\tilde{r}$, $\tilde{r}_A$ in approximating the standard normal under the null hypothesis. Our results show an improvement in the approximation for the second-order method and for $\tilde{r}_A$ with respect to the first-order method, especially in the tails of the distribution. We underline that $\tilde{r}_A$ has been constructed for empirical purposes, that is to reduce the computational effort which is required when operating with $\tilde{r}$. Until now we do not have a theoretical basis to justify this approach, neither an indication of the accuracy of the approximation.

In the following paragraph we report some of simulation studies we realized. They refer to models that are very often used in biological and chemical sciences and in pharmacokinetics.

3.2.1 Simulation studies

Simple logistic models are widely used in bioassays, in particular for the analysis of growth curve data. They relate the response y to an input x via the sigmoidal or “S-shaped” function

$$y = \frac{\phi_1}{1 + \exp\{(\phi_2 - x)/\phi_3\}} + \varepsilon. \quad (15)$$

If the scale parameter ϕ_3 is positive, then ϕ_1 represents the horizontal asymptote as $x \rightarrow \infty$, while 0 is the horizontal asymptote as $x \rightarrow -\infty$. If $\phi_3 < 0$, the roles of ϕ_1 and 0 are reversed. The parameter ϕ_2 is the inflection point of the curve; at this value the expectation of the response is $\phi_1/2$. The simulation study presented here considers $m = 6$ groups with $n_i = 7$ observations each. The following assumptions were made for the group-specific parameters of model (15)

$$\phi_{i1} = \beta_1 \mathbf{1} + b_i \mathbf{1} + \beta_2 v_i, \quad \phi_2 = \beta_3, \quad \phi_3 = \beta_4, \quad (16)$$

with $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)^\top = (200, 0, 750, 130)^\top$. The variable v_i is a 0/1 indicator variable, distinguishing, say, between two different types of treatment, and $\mathbf{1}$ is an n_i -dimensional vector of 1's. The random effect $b_i \sim N(0, 1)$ on ϕ_1 accounts for differences in the asymptotic level among subjects. The within-group errors, ε_{ij} , follow the standard normal distribution. Interest focuses on $\beta_2 = 0$, that is we want to assess whether there is any difference between the two treatments with respect to ϕ_1 . Figure 2 reports the normal Q-Q plots of r , $\tilde{r}$ and $\tilde{r}_A$ obtained from a simulation of $R = 1000$ replicates. As we can see, the normal approximation does not work well for r especially in the tails of the distribution. Instead, it exhibits a satisfactory behaviour for the second-order statistic $\tilde{r}$ and for $\tilde{r}_A$. Table 2 reports the estimated coverage of confidence intervals of different sizes. The statistic r underestimates the nominal level, while $\tilde{r}$ and $\tilde{r}_A$ work well.

The random effect occurs linearly in the model, therefore we can compute the exact form of the log-likelihood. Simulation results for r , $\tilde{r}$ and $\tilde{r}_A$ based on the exact log-likelihood are very similar to the results which are obtained using Lindstrom and Bates' approximation and hence are not displayed.

As a second example, we consider a simple logistic model with the same structure as (15), where now we suppose that the inflection points can vary among the groups of observations. The group-specific parameters of model (15) now are

$$\phi_{i1} = \beta_1 \mathbf{1} + \beta_2 v_i, \quad \phi_2 = \beta_3 + b_i \mathbf{1}, \quad \phi_3 = \beta_4. \quad (17)$$

As before, interest focuses on β_2 . In the simulation, parameters values for β, b_i and ε_i are the same as those used in the previous simulation study.

Figure 3 reports the normal Q-Q plots for r , $\tilde{r}$ and $\tilde{r}_A$ obtained from a simulation of $R = 1000$ replicates. The normal approximation does not work well for r , in particular in the left tail of

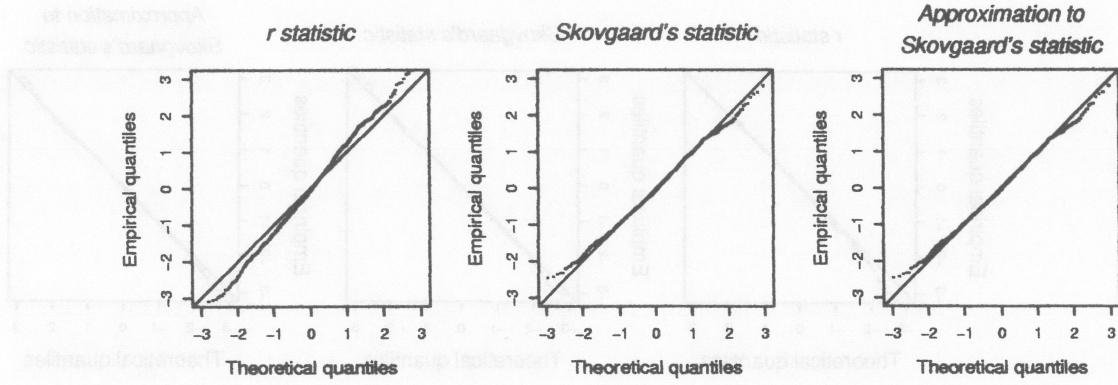


Figure 2: Normal Q-Q plots of r (left), $\tilde{r}$ (middle) and $\tilde{r}_A$ (right) based upon $R = 1000$ replicates of a simulation for testing $\beta_2 = 0$ in model (15) with parameter specification (16).

Coverage	Statistic	
	r	$\tilde{r}, \tilde{r}_A$
0.80	0.706 (0.678,0.734)	0.813 (0.789,0.837)
0.85	0.756 (0.730,0.782)	0.872 (0.852,0.892)
0.90	0.815 (0.791,0.839)	0.923 (0.907,0.939)
0.95	0.909 (0.891,0.927)	0.967 (0.955,0.979)
0.99	0.972 (0.962,0.982)	0.996 (0.992,1.000)

Table 2: Estimated coverage of 80%, 85%, 90%, 95% and 99% confidence intervals for the fixed effect β_2 in model (15) with parameter specification (16), obtained from $R = 1000$ replicates of a simulation.

the distribution. The statistics $\tilde{r}$ and $\tilde{r}_A$ perform better. Table 3 reports the estimated coverage of confidence intervals of different sizes. As before, r tends to underestimate the nominal level, while $\tilde{r}$ and $\tilde{r}_A$ have a satisfactory behaviour.

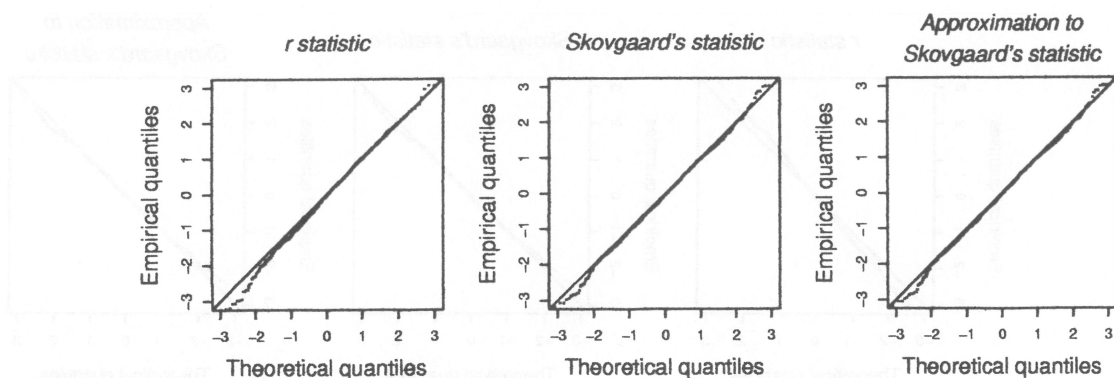


Figure 3: Normal Q-Q plots of r (left), $\tilde{r}$ (middle) and $\tilde{r}_A$ (right) based upon $R = 1000$ replicates of a simulation for testing $\beta_2 = 0$ in model (15) with parameter specification (17).

	Statistic	
Coverage	r	$\tilde{r}, \tilde{r}_A$
0.80	0.775 (0.751,0.799)	0.807 (0.787,0.827)
0.85	0.825 (0.805,0.845)	0.851 (0.833,0.869)
0.90	0.880 (0.860,0.900)	0.898 (0.882,0.914)
0.95	0.934 (0.920,0.948)	0.948 (0.936,0.960)
0.99	0.981 (0.973,0.989)	0.985 (0.979,0.991)

Table 3: Estimated coverage of 80%, 85%, 90%, 95% and 99% confidence intervals for the fixed effect β_2 in model (15) with parameter specification (17), obtained from $R = 1000$ replicates of a simulation.

3.3 Software

Simulation studies for nonlinear mixed-effects models were performed using a series of procedures in R, version 1.3.0 for Microsoft Windows. They constitute an extension of the procedures collected in the `lme_hoa` package. This increase in algebraic complexity with respect to LME models causes a larger computational burden. As in the linear case, we referred to `nlme` library of Pinheiro and Bates (2000) for parameters estimation. For details, see Guolo (2001, chapter 6).

4 Final remarks

Results of the application of Skovgaard's statistic $\tilde{r}$ to linear mixed-effects models have empirically shown that the improvement of the standard normal approximation for the null distribution is remarkable. The accuracy in the approximation is satisfactory most of all in the tails of the distribution.

Also for the nonlinear case, simulation results were satisfactory. However, some remarks are needed. First, the calculation of $\tilde{r}$ relies on Lindstrom and Bates' approximation. This means that the log-likelihood and $\tilde{r}$ are not calculated using the true MLEs of parameters, even though some results in the literature indicate that approximate MLEs are very close to the exact ones (Pinheiro and Bates, 1995). Moreover, the calculation of Skovgaard's statistic is more complicated for NLME models than for LME models. Under this respect, the approximation $\tilde{r}_A$ is a simple but effective alternative. As we have already remarked, this statistic has been proposed here for strictly empirical purposes. It has shown satisfactory results, but these are restricted to the simulation experiments we realized. Hence, a further study of the characteristics of $\tilde{r}_A$, especially the theoretical ones, seems to be necessary.

Acknowledgements

The topics discussed in this paper have been developed as a part of my graduate thesis submitted to the Faculty of Statistics, University of Padova, under the supervision of Alessandra Salvan and Alessandra R. Brazzale.

References

- Barndorff-Nielsen, O.E. (1986). Inference of full and partial parameters based on the standardized signed log likelihood ratio. *Biometrika*, **73**, 307-322.
- Davidian, M. and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall, London.
- Guolo, A. (2001). *Verosimiglianza e metodi asintotici in modelli con effetti misti*. Tesi di laurea, Facoltà di Scienze Statistiche ed Economiche, Università di Padova.
- Jenrich, R.I. and Schluter, M.D. (1986). Unbalanced repeated measures models with structured covariance matrices. *Biometrics*, **42**, 805-820.
- Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 963-974.

- Lindstrom, M.J. and Bates, D.M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated measures data. *Journal of the American Statistical Association*, **83**, 1014-1022.
- Lindstrom, M.J. and Bates, D.M. (1990). Nonlinear mixed-effects models for repeated measures data. *Biometrics*, **46**, 673-687.
- Lyons, B. and Peters, D. (2000). Applying Skovgaard's modified directed likelihood statistic to mixed linear models. *Journal of Statistical Computation and Simulation*, **65**, 225-242.
- Pinheiro, J.C. and Bates, D.M. (1995). Approximations to the log-likelihood function in nonlinear mixed-effects models. *Journal of Computational and Graphical Statistics*, **4**(1), 12-35.
- Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag, New York.
- Robinson, G.K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science*, **6**, 15-51.
- Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components*. John Wiley & Sons, New York.
- Severini, T.A. (2000). *Likelihood Methods in Statistics*. Oxford University Press, Oxford.
- Skovgaard, I.M. (1996). An explicit large-deviation approximation to one-parameter tests. *Bernoulli*, **2**, 145-165.